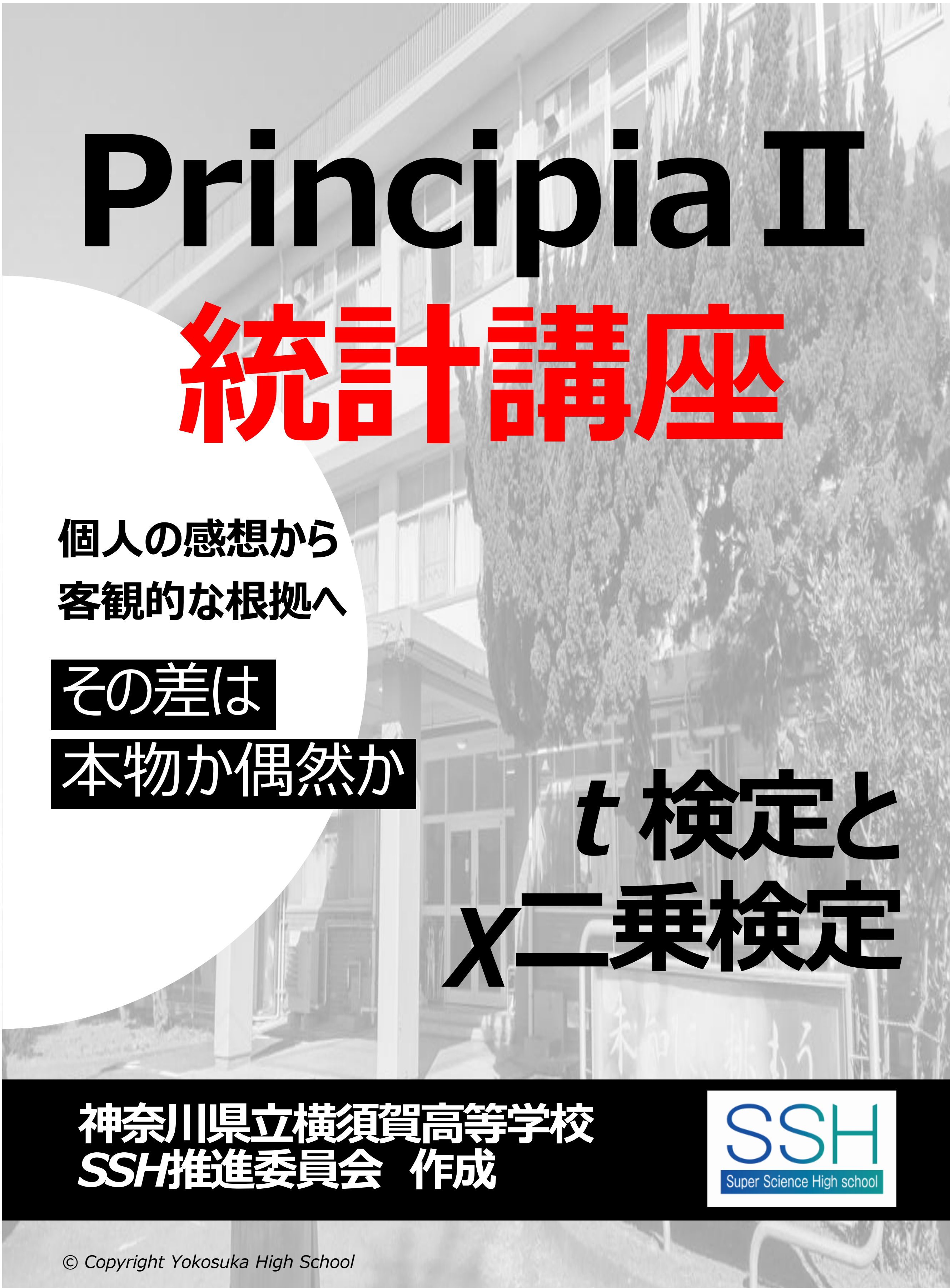


Principia Ⅱ

統計講座

個人の感想から
客観的な根拠へ

その差は
本物か偶然か

t 検定と
χ二乗検定

神奈川県立横須賀高等学校
SSH推進委員会　作成

Super Science High school

© Copyright Yokosuka High School

PrincipiaⅡ　統計講座

【第１章】はじめに

１．なぜ統計処理が必要なのか？

皆さんは探究学習で、「A よりも B の方が良い結果が出た」「アンケートの結果、〇〇という傾向が見られた」といった結論を出そうとしているはずです。しかし、そこで立ち止まって考えてみてください。

その「差」や「傾向」は、本当に意味のあるものですか？それとも、たまたま起こった「誤差」ですか？

２．「たまたま」に騙されないために

例えば、ある薬を飲んだ５人のうち３人に効果があったとします。これだけで「この薬は効果がある！」と言い切れるでしょうか？たった５人なら、偶然体調が良かった人が集まっただけかもしれません。統計処理（仮説検定）は、目の前にあるデータの差が「偶然のいたずら」なのか、それとも「意味のある必然」なのかを、数学的なルールに則って白黒つけるためのツールです。

３．主観を捨てて、誰もが納得する「証拠」を作る

「グラフを見れば一目瞭然だ」と思うかもしれません。しかし、人によって「これくらいなら差があると言える」という基準はバラバラです。「10%の差があれば十分だ」と言う人「いや、30%は違わないと信じない」と言う人これでは、せっかくの発見も説得力を持ちません。統計学は、「誰が判断しても同じ結論になる共通の物差し」を与えてくれます。$p < 0.05\%$（5%未満の確率でしか起こらない珍しいこと）といった世界共通のルールを使うことで、あなたの探究結果は「個人の感想」から「客観的な事実」へと進化します。

４．複雑な世の中から「本質」を見抜く

世の中のデータには必ず「バラツキ」があります。

・同じ種をまいても、育つ高さはバラバラです。

・同じテストを受けても、点数は一人ひとり違います。

このノイズ（バラツキ）だらけのデータの中から、「本当に影響を与えている原因は何なのか？」をあぶり出すのが、今回学ぶ t 検定や χ 二乗検定（カイニジョウ検定）といった手法です。

５．この講座のゴール

難しい公式を暗記することではありません。自分の出したデータに対して、「これは偶然ではないと言えるか？」という問いに、自信を持って答えられるようになることです。統計という「最強の武器」を手に入れて、あなたの探究を「根拠のある科学」へと変えていきましょう。

==

【第２章】仮説検定：その差は「本物」か「偶然」か？

仮説検定とは、「ある仮説が正しいと言えるかどうかを、確率を使って判断する手続き」のことです。１年生の数学で学んだと思いますが、少し復習しましょう。

1．背理法のような考え方

仮説検定は、「言いたいことの反対（帰無仮説）を、あえて仮定」します。「背理法」のような考え方です。

- **帰無仮説（H_0）**：「差はない」「効果はない」「偶然である」という仮説
- **対立仮説（H_1）**：「差がある」「効果がある」「偶然ではない」という仮説（あなたが証明したいこと）

※なぜわざわざ「差がない」と仮定するのか？

「差がある」ことを証明するのは難しいですが、「『差がない』と考えるには無理がある（ありえない！）」と矛盾を突く方が、数学的に証明しやすいからです。

２．判断のプロセス（4 つのステップ）

例えば、「新しく開発した肥料で植物の背が高くなるか」を調べる場合で考えてみましょう。

①仮説を立てる

帰無仮説：「肥料をあげてもあげなくても、背の高さは**変わらない。**」

②計算して確率を出す（p 値）

実験データから、「もし肥料に効果がないとしたら、今回の結果が起こる確率はどれくらいか？」を計算します。この確率を **p 値**と呼びます。

③基準と比べる（有意水準）

あらかじめ決めておいたボーダーライン（一般的には **5%** ）と p 値を比べます。

④結論を出す

$p < 0.05$ のとき

5%より低い確率＝「めったに起きないことが起きた！」と判断し、帰無仮説を捨てます（**棄却**）。
結果、「差がある（有意である）」と言えます。

$p \geqq 0.05$ のとき

「偶然の範囲内」と判断し、帰無仮説を捨てられません。結果、「差があるとは言えない」となります。

３．「有意差あり」が意味すること・しないこと

検定の結果、「有意差あり（$p < 0.05$）」となった時に勘違いしてはいけないポイントが 2 つあります。

①「絶対に正しい」わけではない

たまたま珍しいことが起きただけの可能性も 5%未満ですが残っています。あくまで「確率的に正しいらしい」という判断です。

②「差が大きい」とは限らない

統計的に「差がある」と言えても、その差がごくわずか（例えば身長が 0.1mm だけ高いなど）で、実用上の意味がない場合もあります。

【ワーク】仮説を立ててみよう

あなたの探究テーマで、以下の空欄を埋めてみてください。

- **あなたが確かめたいこと（対立仮説）：**＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿（例：筋トレをするとテストの点数が上がる）
- **あえて否定した仮説　　（帰無仮説）：**＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿（例：筋トレをしてもテストの点数は上がらない）

まとめ

「p 値が 0.05 より小さければ、それは『偶然』ではない！」

次は、具体的なデータの種類（平均値なのか、割合なのか）に合わせて、どの検定（ t 検定や χ 二乗検定）を使えばいいかを学びます。

==

【第３章】t 検定と χ 二乗検定

t 検定と χ 二乗検定は、扱うデータの「種類」が全く異なります。

一言でいうと、t 検定は「平均値（数値）」を比べ、χ 二乗検定は「割合（人数・回数）」を比べます。

１． t 検定（平均値の差を調べる）

身長、体重、テストの点数、気温など、「連続的な数値」の平均に差があるかどうかを判定します。

- **データの種類：** 数値（単位があるものが多い）
  - 「A 組と B 組で数学のテストの**平均点**に差があるか？」
  - 「新薬を飲む前と後で、**血圧**が下がったか？」
- **考え方：** ２つのグループの平均がどれくらい離れているかを見て判断します。

２． χ 二乗検定（割合の差・関連を調べる）

合格・不合格、YES・NO など、「カテゴリー（項目）」ごとの人数や割合に偏りがあるかを判定します。

- **データの種類：** カウントデータ（人数、個数、回数）
  - 「A 組と B 組で、テストの**合格率**に差があるか？」
  - 「性別によって、支持する政党の**割合**に偏りがあるか？」
- **考え方：** 「もし差がないとしたら、これくらいの人数になるはずだ」という理論値（期待値）に対し、実際の結果がどれだけズレているかを見ます。

どっちを使えばいい？（比較表）

t 検定	χ 二乗検定
平均値に差があるか	割合・関連に差があるか
点数、時間、長さ、重さ	男女、血液型、良品/不良品
平均と標準偏差（バラツキ）	実測値と期待値のズレ
A 組の方が平均 5 点高い	男性の方が、購入率が 20%高い

豆知識： もし「A 組と B 組で、テストが 80 点以上の人の**割合**に差があるか」を調べたいなら、元は点数（数値）でも「80 点以上か未満か」というカテゴリーに変えているので、χ 二乗検定を使います。
分析しようとしているデータをどの検定で検定したらいいか分からなかったら、AI に聞いてみましょう。ただし、検定に関する基本的知識がなければ、AI が正しい提案をしているか判断ができないので、基本的知識を持ったうえで AI を活用する必要があります。

練習問題：調査したい内容①～⑤は「t 検定と χ 二乗検定」どちらを使って検定すると良いか。

調査したい内容	t 検定 or χ 二乗検定
①A 小学校と B 小学校の児童の平均身長に差があるか？	
②広告 A と広告 B で、クリックした人の割合に差があるか？	
③新薬を投与したグループと偽薬のグループで、血圧の低下量に差があるか？	
④男女で、好きなスマホの機種（iPhone か Android か）に偏りがあるか？	
⑤昨年と今年で、1 人あたりの年間平均読書数に変化があったか？	

解説

==

３．両側検定（Two-tailed Test）

「とにかく差があるかどうか」を調べたいときに使います。

- **考え方：**A と B を比較して、「A の方が高い」場合も「A の方が低い」場合も、どちらも「差がある」とみなしてチェックします。
- **いつ使う？：**２つのクラスの平均点に「違いがあるか」だけを知りたいとき、など。
- **特徴：**判定が少し厳しくなります。5%の棄却域を、右端に 2.5%、左端に 2.5%と分けて配置します。

４．片側検定（One-tailed Test）

「あらかじめ方向が決まっている」ときに使います。

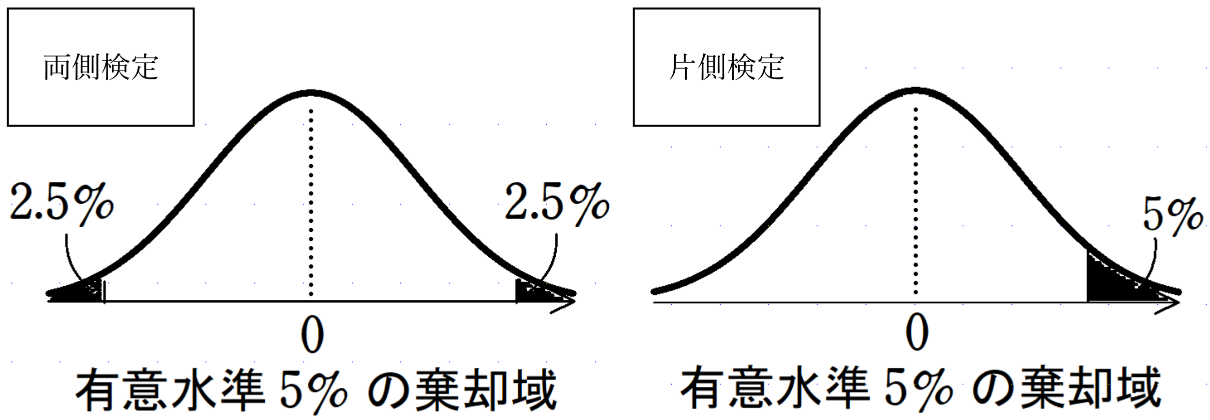
- **考え方：**「A の方が高い（低い）」ことだけチェックし、逆の結果が出ても「差がない」とみなします。
- **いつ使う？：**「この肥料をあげれば、絶対に背が高くなる（低くなることはない）」と断言できるとき。
- **特徴：**5%の棄却域を片方の端っこだけに集中させるため、両側検定よりも有意差が出やすくなります。

５．どっちを使えばいいの？

項目	両側検定	片側検定
目的	違い（変化）があるか知りたい	特定の方向に差があるか知りたい
厳格さ	厳しい（ミスが少ない）	甘め（差が出やすい）
探究での推奨	基本はこちら	特殊な理由がある時だけ

６．Principia では基本的に「両側検定」を使おう

片側検定は「差が出やすい」というメリットがありますが、「都合の良いデータだけを拾おうとしている」と疑われるリスクがあります。例えば、「新しい勉強法で点数が上がるはずだ！」と信じて片側検定をしたのに、実際には点数がガクッと下がってしまった場合、片側検定のルールではその結果を無視することになってしまいます。これでは科学的な分析とは言えません。論文や発表の場で突っ込まれないためには、「特に理由がない限り、最も厳しい基準である両側検定を採用しました」と言えるようにしておくのが一番安全です。



==

【第４章】Excel の「データ分析」機能を使って「t 検定」をやってみよう

次は、具体的なデータを用いて t 検定をやってみましょう。それでは、surface 等の PC を起動し、Excel を開いてください。

Excel の「データ分析」機能を活用してみよう。

Excel を起動する。

① 上部にある「ファイル」をクリック

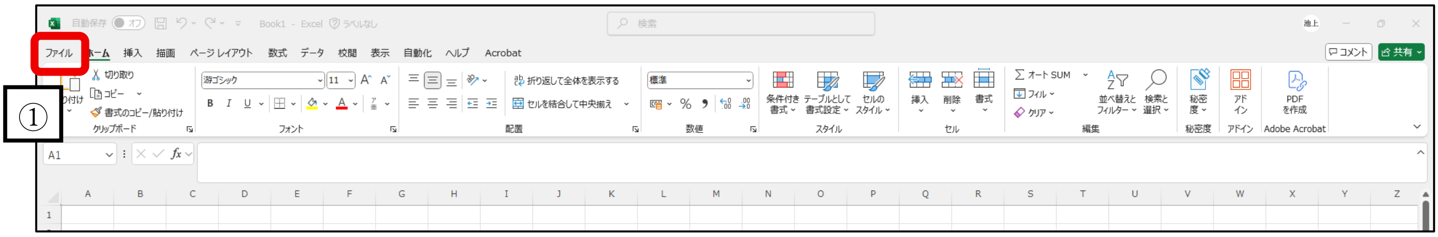
② 開いたウィンドウの左下の「オプション」をクリック

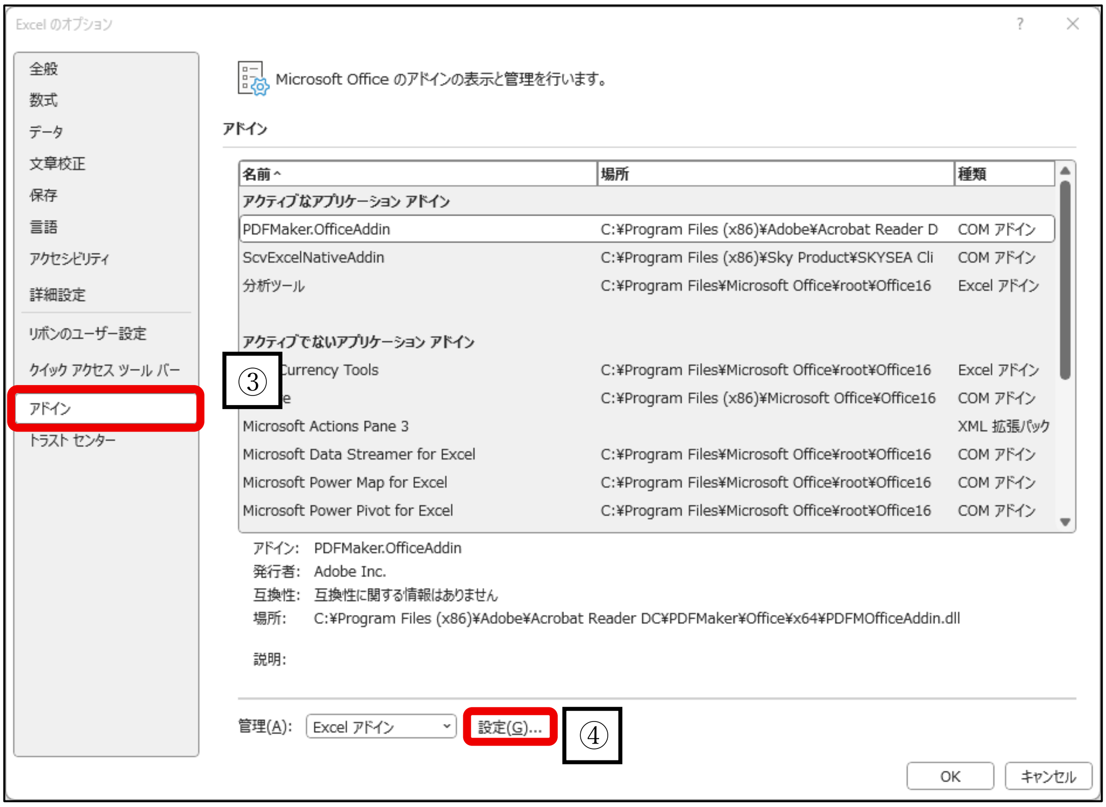
③ オプションウィンドウの「アドイン」をクリック

④ 下部の「設定」をクリック

⑤ 「□分析ツール」にチェックを入れる。(☑分析ツール)

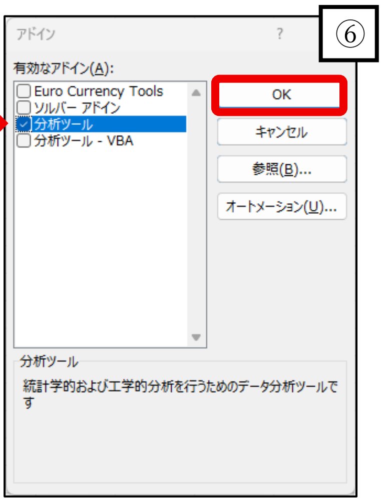
⑥ 「OK」をクリック





チェック

⑤

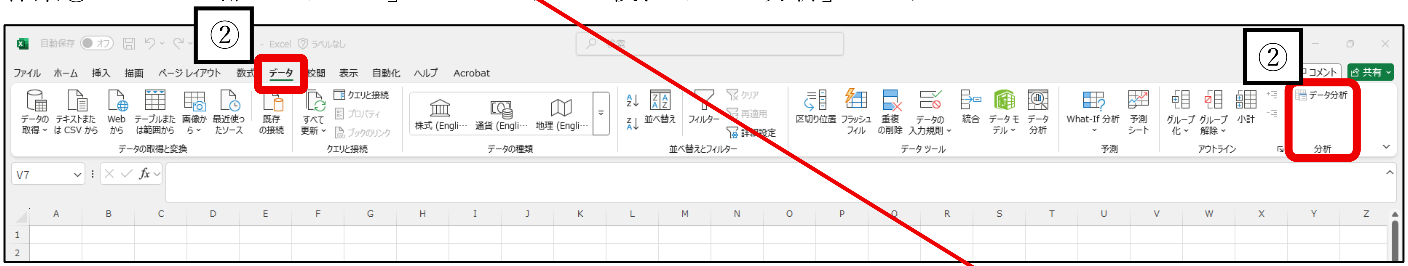


これで「分析ツール」を Excel 上に設定することが出来ました。次は、具体的なデータを使って「分析ツール」を使ってみましょう。

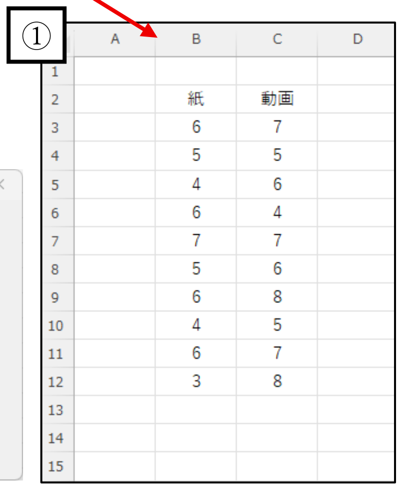
例題　紙で暗記した場合と動画を使って暗記した場合で、暗記しやすさに違いがあるのか検証したい。
10 人の被験者に、紙を使って暗記してもらった場合と、動画を使って暗記した場合で、テストの結果に差があるか、調査した。テストは 10 点満点で、それぞれの被験者の得点は以下の通りである。

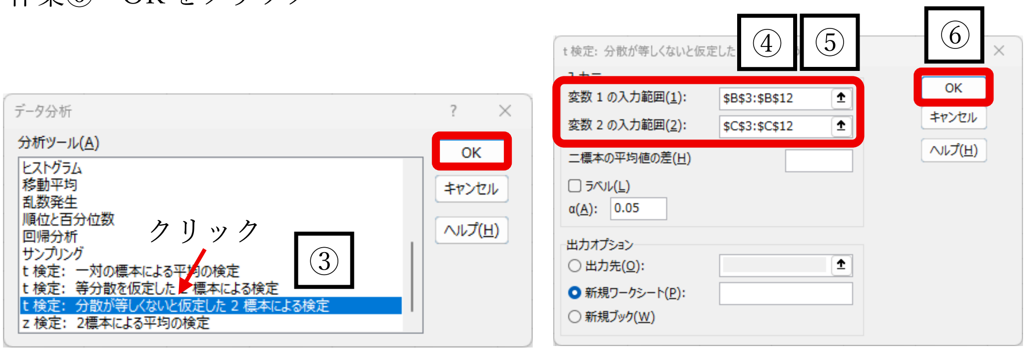
	A さん	B さん	C さん	D さん	E さん	F さん	G さん	H さん	I さん	J さん
紙	6 点	5 点	4 点	6 点	7 点	5 点	6 点	4 点	6 点	3 点
動画	7 点	5 点	6 点	4 点	7 点	6 点	8 点	5 点	7 点	8 点

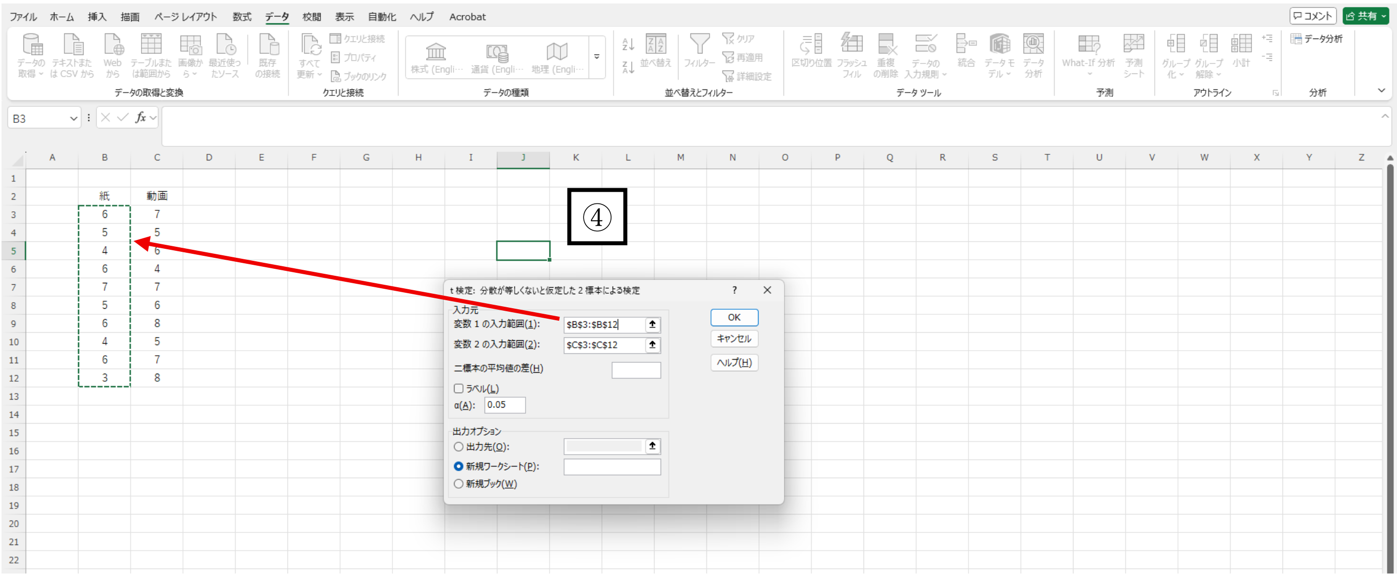
作業①　得点を Excel に入力する。
作業②　Excel 上部の「データ」タブをクリック後、「データ分析」をクリック



作業③　「t 検定：分散が等しくないと仮定した 2 標本による検定」をクリックしたあと、OK をクリック
作業④　変数 1 の入力範囲(1)：をクリックし、入力範囲を選択する。
作業⑤　変数 2 の入力範囲(2)：をクリックし、入力範囲を選択する。
作業⑥　OK をクリック







作業⑦　新規ワークシートに検定結果が表示されるので、以下の表に転記する。

	変数 1	変数 2
平均		
分散		
観測数		
仮説平均との差異		
自由度		
t		
P(T<=t) 片側		
t 境界値 片側		
P(T<=t) 両側		
t 境界値 両側		

作業①「t」値に絶対値をつけてプラスにし、下記表に記入する。小数第三位を四捨五入
作業②「t 境界値　両側」の値を下記表に記入する。
作業③「t」値、「t 境界値　両側」の値について、大小を評価する。

「t」値をプラスにしたもの	不等号	「t 境界値 両側」の値

「P(T<=t) 両側」の値

もし、本当は差がない（帰無仮説が正しい）としたら、今回のようなデータ（あるいはそれ以上に極端なデータ）が得られる確率は約 7%である

❶「t」値 ＜ 「t 境界値 両側」の値　の場合
「統計的に有意な差があるとはいえない（帰無仮説を棄却できない)」という状態です。
つまり「データの差が、偶然の範囲内に収まっている可能性が高い」という判定になります。

❷「t」値 ＞ 「t 境界値 両側」の値　の場合
「統計的に有意な差がある（帰無仮説を棄却できる」という状態です。
つまり「偶然とは考えにくいほどのハッキリとした差がデータに現れている」という判定になります。

※ただし、「棄却できた」からといって、「100%絶対に正しい」わけではありません。有意水準 5%で検定したなら、「5%くらいの確率で、本当は差がないのに、たまたま極端なデータが出てしまって間違った判断をしている可能性」はわずかに残っている、ということは頭の片隅に置いておいてください。

有意水準 5%で検定しています。

次は、具体的なデータを使って、t 検定を行ってみましょう。

姿勢を変えるとテストの点数は何点上がるか？　―「個人の感想」から「科学的な根拠」へ―

STEP 1：実験とデータの収集

クラスの生徒○人を対象に、「正しい姿勢」と「突っ伏した姿勢」の両方でテストを行い、データを集めます。

•目的：姿勢の良し悪しが、脳のパフォーマンス（テストの点数）にどう影響するかを確かめる。

•方法：同一人物が両方の姿勢でテストを受け、その「差（伸び代）」に注目する。

==

STEP 2：t 検定（それは「本物」か「偶然」か？）

平均して○○点上がったとしても、「たまたまその時の運が良かっただけじゃない？」という疑いが残ります。これを科学的に論破するのが t 検定です。

•チェックする数字：p 値（P(T<=t) 両側）

•出た結果：○．○○（○%）

•判定：この数字が 0.05（5%）より小さければ、「姿勢で点数が変わる」という現象は、ほぼ 100%に近い確率で「本物」であると証明されます。偶然でこの結果が出る確率は、○%程度しかありません。

==

STEP 3：母集団の推定（みんなに当てはめると？）

次に、この「○人の結果」を使って、まだ実験していない「全校生徒」や「未来の自分」にどれくらいの効果があるかを予測します。

・使う数字（基本統計量より）：平均値：○点　信頼区間：○点

・解釈：「この○人のデータから推測すると、他の人がやっても、あるいは次におこなっても、95%の確率で
　　○点〜○点の間でスコアアップが期待できる」と言えます。

==

STEP 4：最終結論

実験と統計分析により、正しい姿勢は集中力を高め、スコアを確実に押し上げることが分かりました。その効果は全校生徒に当てはめた場合、平均して○点〜○点の伸び代として現れます。

==

💡 この授業の学びのポイント

１．t 検定は、そのデータが「信じるに値するか」を決める門番。

２．母集団の推定とは、得られた平均値などから、『この区間「○○点〜○○点」は、95%の確率で母集団の平均値を含むはずだ』といったいわゆる「信頼区間」を求めることをいいます。この区間を求めることで、抽出した集団から得られた数値から母集団の数値を推測することが出来ます。

この２つが揃うことで、あなたの主張は「ただの感想」から「誰もが納得する科学的根拠」に変わります。

練習問題１

クラスの 10 人に「正しい姿勢」と「悪い姿勢」で計算してもらいました。その結果について t 検定を実施し、p 値に基づいて有意差の有無を判定し、この結果から言えることを班で共有しなさい。

出席番号	正しい姿勢 A	悪い姿勢 B	差　A-B
1	38	35	
2	32	34	
3	45	40	
4	28	30	
5	40	36	
6	35	32	
7	42	43	
8	30	28	
9	48	44	
10	33	31	
平均			

t 検定の P 値

練習問題２

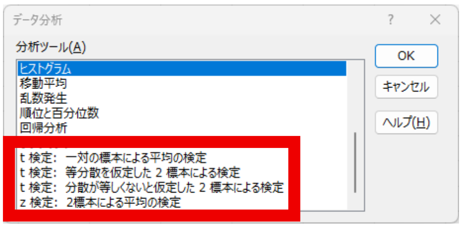
練習問題１のデータを４回コピーして、40 人分のデータにしなさい。そのデータを使って t 検定を実施し、p 値に基づいて有意差の有無を判定し、この結果から言えることを班で共有しなさい。

【サンプル数について】

例えば「クジラ（大きな差）を捕まえるなら、網目が粗くても（10 人）でも引っかかる。でも、メダカ（小さな差）を捕まえたいなら、網目が細かい網（100 人）を仕掛けないと、網の目をすり抜けて『有意差なし』になってしまう。

【検定の使い分けについて】

Excel の「データ分析」機能には、t 検定の種類がいくつかある。これらの違いは以下の通りである。



<table>
<tr><th colspan="2">検定の名前</th><th>どんな時に使う？</th><th>データの関係</th></tr>
<tr><td rowspan="3">t 検定</td><td>一対の標本による平均の検定
【通称】対応のある t 検定</td><td>同じ人の変化を見たいとき。</td><td>A さんと A さん</td></tr>
<tr><td>等分散を仮定した２標本による検定
【通称】スチューデントの t 検定</td><td>別の人で、バラつき（分散）が同じとき。</td><td>A さんたちと B さんたち</td></tr>
<tr><td>分散が等しくないと仮定した２標本による検定
【通称】ウェルチの t 検定</td><td>別の人で、バラつき（分散）が違うとき。（現代の統計学では、等分散かどうかに関わらず、最初から「分散が等しくない」方を使う方が主流です。</td><td>A さんたちと B さんたち</td></tr>
<tr><td>Z 検定</td><td>２標本による平均の検定</td><td>母集団の分散が分かっているなど、神様レベルで背景を知っているとき。数 B の教科書で学ぶ統計的な推測では、なんと Z 検定を行う。</td><td>特殊なケース（ほぼ使わない）</td></tr>
<tr><td colspan="2">χ 二乗検定</td><td>属性（性別など）によって「割合」や「好み」に差があるか見たいとき。</td><td>A さんたちと B さんたち</td></tr>
</table>

練習問題 3

クラスの 10 人を対象に、「正しい姿勢」と「悪い姿勢」のそれぞれで計算テストを行ったところ、以下の結果が得られた。この結果をもとに、学校全体の生徒 800 人が同じ条件でテストを受けた場合の、平均点の 95%信頼区間をそれぞれ求めよう。

出席番号	正しい姿勢 A	悪い姿勢 B	差　A-B
1	38	35	
2	32	34	
3	45	40	
4	28	30	
5	40	36	
6	35	32	
7	42	43	
8	30	28	
9	48	44	
10	33	31	
平均			

この 10 人における「正しい姿勢」の平均点は（　　　　）点、標準偏差は（　　　　）点なので、誤差は（　　　　）点となり、母集団の平均点の 95%の信頼区間は（　　　　　　　　　）となります。

Excel の関数で標準偏差を求めるとき、以下に気を付けよう。
・STDEV.P：計算に使うデータが，全データのとき　　　　（高 1 で習った標準偏差）
・STDEV.S：計算に使うデータが，全データの一部のとき　（高 1 で習った標準偏差と分母が違う）

STDEV.P の場合　**標準偏差**　$s=\sqrt{\frac{1}{n}\{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots\cdots+(x_n-\overline{x})^2\}}$

STDEV.S の場合　**標準偏差**　$s=\sqrt{\frac{1}{n-1}\{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots\cdots+(x_n-\overline{x})^2\}}$

少ないサンプル数から得られた標準偏差は少し小さめに出てしまう傾向があるので、$n-1$で割ることにより標準偏差を少し大きくしています。これは高校数学では習いません。この計算方法が「STDEV.S」なのです。今回のように全体の一部のデータの標準偏差を求めるときは、$n-1$で割る STDEV.S を使う方が良いでしょう。ちなみに、STDEV.P の P は Population（母集団)、STDEV.S の S は Sample（標本）のことです。

高校数学では、95%の信頼区間における誤差は $1.96\cdot\frac{\sigma}{\sqrt{n}}$ と計算します。（σ はシグマと読み、標準偏差のこと）しかし、今回はデータ数が少ないので、高校数学で使う理論値 1.96（データ数∞）を使うより、t 分布（サンプル数によりそれぞれ決められた値、サンプル数 10 なら自由度 9 なので 2.26）に従って決められた数値を使い誤差を出す方が、より良いでしょう。この t 分布は Excel のデータ分析が自動で使って計算してくれます。

結果、「正しい姿勢」の母集団における 95%の信頼区間は「32.3 点～41.8 点」となるのです。この区間は、95%の確率で母集団平均を含むわけですが、これだと幅がありすぎますよね。なので、**サンプル数を大きくするとより狭い範囲の信頼区間を求めることが出来る**わけです。例えば、サンプル数を 4 倍の 40 人にすると、分母が$\sqrt{10}$ から$\sqrt{40}$ に変化するので、信頼区間の幅は理論上半分になります。なお、高校数学では常に 1.96 を使っているので、単純に誤差を半分にして平均に±すれば 95%の信頼区間が求められますが、今回は t 分布を使っているので、データが 40 人に増えた場合サンプル数に応じた数値が 2.26→2.02（データが増えたので 1.96 に近づきましたね）に変わります。つまり、皆さんは Excel の計算に任せましょう。

（参考）t 分布　東京都立大学 学術情報基盤センターHP より抜粋

自由度	確率95%	確率99%	自由度	確率95%	確率99%
1	12.706	63.657	18	2.101	2.878
2	4.303	9.925	19	2.093	2.861
3	3.182	5.841	20	2.086	2.845
4	2.776	4.604	21	2.080	2.831
5	2.571	4.032	22	2.074	2.819
6	2.447	3.707	23	2.069	2.807
7	2.365	3.499	24	2.064	2.797
8	2.306	3.355	25	2.060	2.787
9	2.262	3.250	26	2.056	2.779
10	2.228	3.169	27	2.052	2.771
11	2.201	3.106	28	2.048	2.763
12	2.179	3.055	29	2.045	2.756
13	2.160	3.012	30	2.042	2.750
14	2.145	2.977	40	2.021	2.704
15	2.131	2.947	60	2.000	2.660
16	2.120	2.921	120	1.980	2.617
17	2.110	2.898	∞	1.960	2.576

t分布表

Excel を使って 95%の信頼区間を求めてみよう。

作業①　得点を Excel に入力する。

作業②　Excel 上部の「データ」タブをクリック後、「データ分析」をクリック

作業③　データ分析ツールのうち、基本統計量を選んで OK をクリック

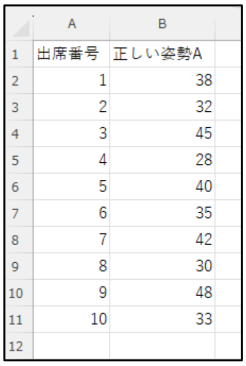
作業④　入力範囲を選んで、統計情報と平均の信頼度の出力にチェックを入れ OK をクリック

作業⑤　新規ワークシートに出てきたデータの内容を以下に転記する。

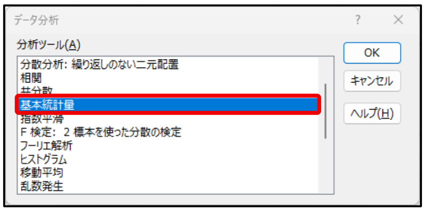
平均	
信頼度（95.0%）の値	
平均＋信頼度（95.0%）の値	
平均－信頼度（95.0%）の値	
95%の信頼区間	

②

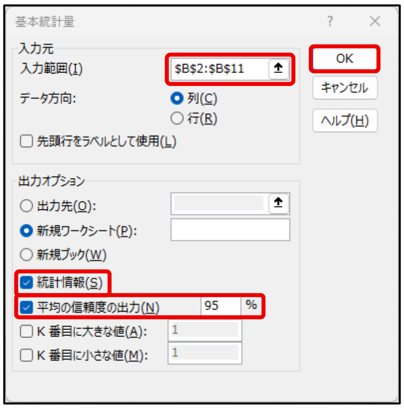
①



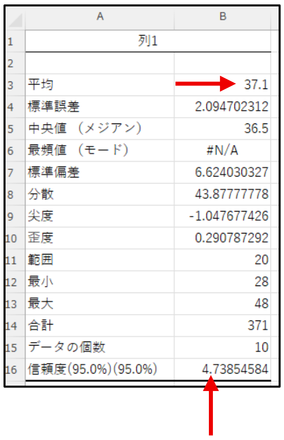
③



④



⑤



【第５章】カイ二乗検定：割合や好みの「偏り」を暴く

前章では、テストの点数や身長などの「数値の平均」を比べる t 検定を学びました。今回は、「賛成・反対」「合格・不合格」のように、**カテゴリーごとの人数（頻度）に差があるかどうかを調べる χ 二乗検定**を学びます。

１．χ 二乗検定とは？

一言でいうと、「実測値」と「期待値」のズレを調べる検定です。

- 実測値： 実際に実験やアンケートで得られた人数。
- 期待値： もし「差がない」としたら、理論上こうなるはずだという人数。

考え方の例：コイン投げ

コインを 100 回投げて、表が 60 回、裏が 40 回出たとします。

- 期待値： 差がないなら「表 50 回、裏 50 回」のはず。
- 実測値： 実際は「表 60 回、裏 40 回」。

この「10 回のズレ」は、たまたま起こる範囲内（誤差）か、それともコインに仕掛けがある（有意差）のか？

…これを判定するのがカイ二乗検定です。

２．独立性の検定（２つの項目の関係性を調べる）

探究学習で最もよく使われるのが、この「独立性の検定」です。

例：「性別」によって「朝食の摂取習慣」に違いがあるか？

（「性別」と「朝食」という２つの要素が、お互いにバラバラか、それとも関係があるかを調べます）

【実践】Excel を使ってカイ二乗検定をやってみよう

Excel には t 検定のような便利な「データ分析ボタン」がカイ二乗検定にはありません。

そのため、「CHISQ.TEST（カイ・スクエア・テスト）関数」を使います。

例題：スマートフォンの OS 利用率調査

ある高校で、男子 50 人、女子 50 人に「使っているスマホは iPhone か Android か」を調査したところ、以下の結果になった。性別によって、選ぶスマホに違いがあると言えるか（有意差があるか）検証せよ。

性別	iPhone	Android	合計
男子	30	20	50
女子	40	10	50
合計	70	30	100

	A	B	C	D
1	性別	iPhone	Android	合計
2	男子	30	20	50
3	女子	40	10	50
4	合計	70	30	100
5				
6		iPhone	Android	
7	男子			
8	女子			
9				
10	p値			
11				

手順①：Excel に「実測値」の表を作る

手順②：「期待値」の表を作る

　期待値の計算式：（行の合計 × 列の合計）÷ 全体の合計

　例：男子の iPhone 期待値：(50×70)÷100 =

　例：女子の Android 期待値：(50×30) ÷100 =

手順③：CHISQ.TEST 関数で p 値を出す

　適当なセルに以下の関数を入力します。

　=CHISQ.TEST(実測値の範囲，期待値の範囲)

手順④：結果を判定する

　計算の結果、p＝　　　　となりました。

ちなみに、高校数学で習う期待値は、「1 回試行したときに、平均してどれくらいの値が得られるか」という予測値です。しかし、**統計学（検定）の期待値は、「もしグループ間に差がないとしたら、理論上ここには何人いるはずか」という人数（頻度）**です。